STATE-OF-THE-ART EVENT GENERATORS VIA NEW TECHNIQUES AND TECHNOLOGIES*

*AN ENTIRELY SUBJECTIVE ACCOUNT

SNOWMASS COMMUNITY PLANNING MEETING – 7TH OCTOBER 2020
ADVANCES IN EVENT GENERATION AND DETECTOR SIMULATION

ENRICO BOTHMANN

GEORG-AUGUST-UNIVERSITÄT GÖTTINGEN

Institut für Theoretische Physik

INTRODUCTION

motivation

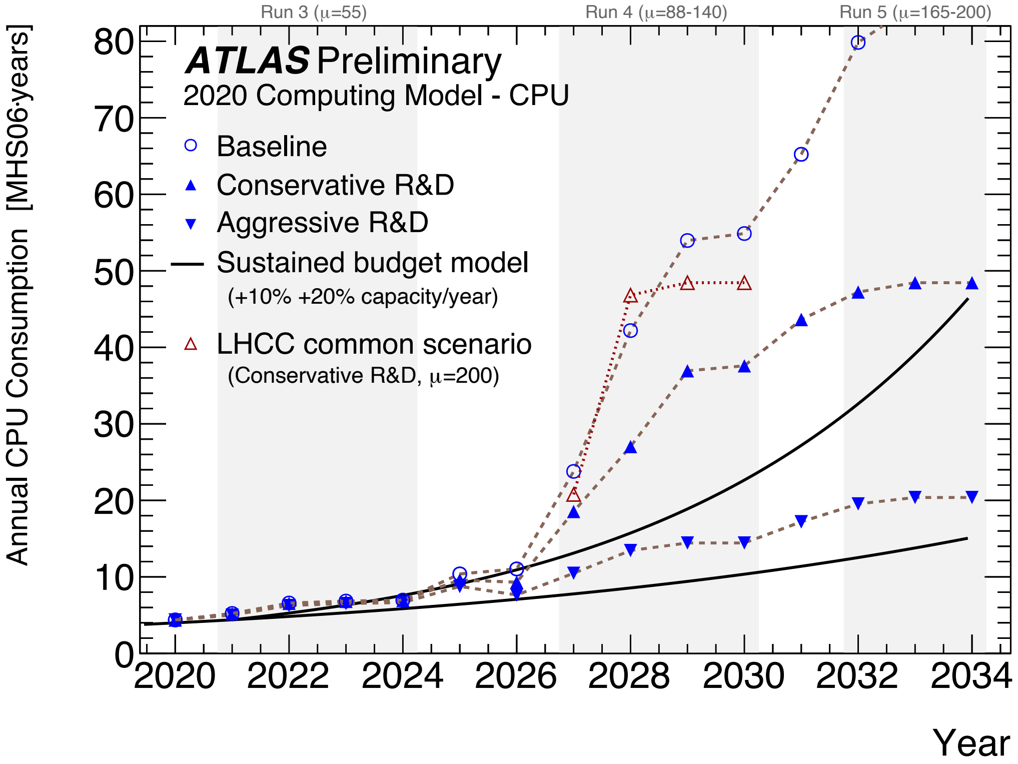
→ Josh' talk

do we want to make compromises in precision or have physics analyses limited by MC stats?

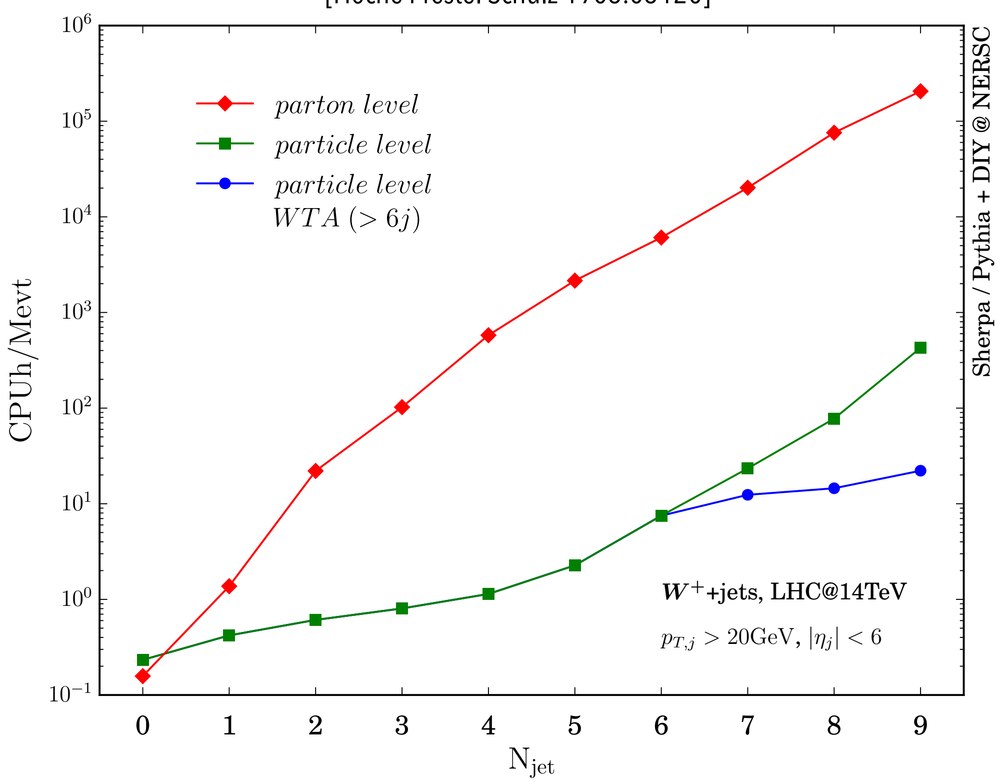
going for the bottleneck: parton-level event generation

1. less ME evaluations by
    - storing parton-level events/grids
    - doing a-posteriori/on-the-fly reweighting
    - improving sampling efficiency → slide #2
2. more throughput: harness parallel computing power now that Moore's law for CPU has stopped working for us (there is no free lunch *any longer*) → slide #3

[https://twiki.cern.ch/twiki/bin/view/AtlasPublic/ComputingandSoftwarePublicResults]



[Höche Prestel Schulz 1905.05120]

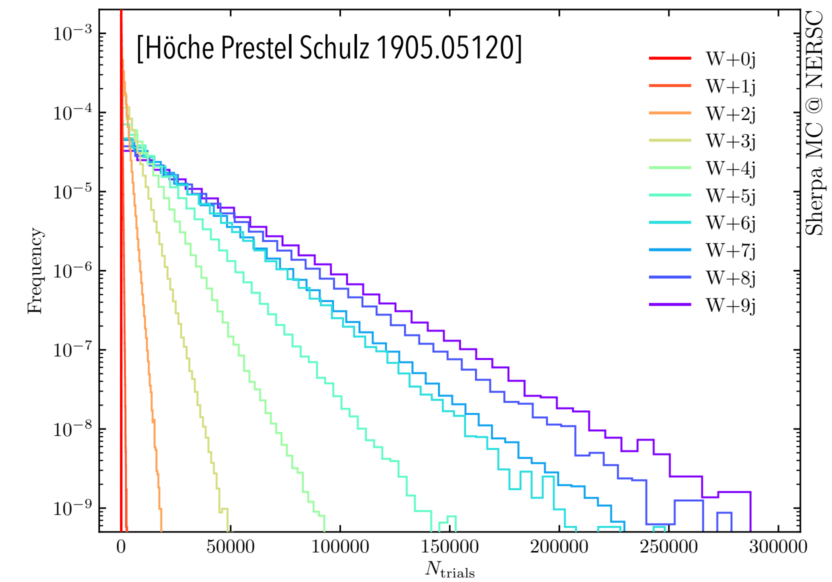


REDUCE ME EVALUATIONS

the issue

unweighting efficiency $\langle w \rangle / w_{\text{max}}$ degrades with process complexity

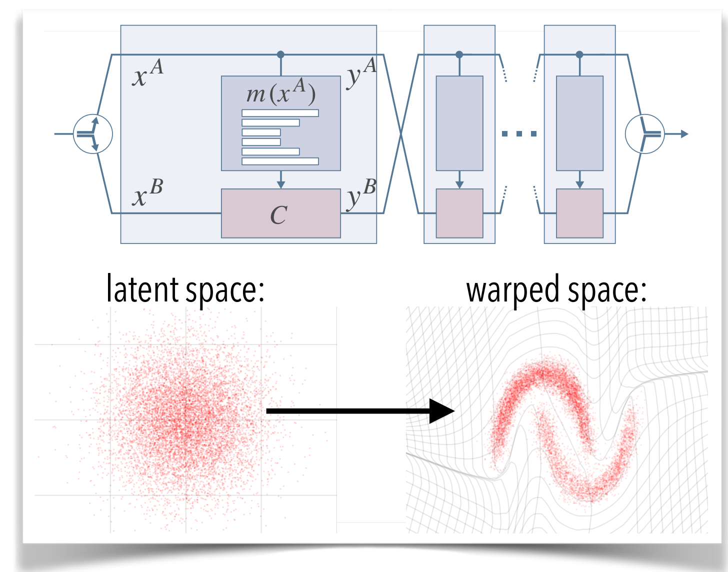
need many parton-level trial events (i.e. ME evaluations) to generate single unweighted event due to wide weight distribution of trial events



the idea

replace 1D ML algorithms (VEGAS) with more flexible DNN-powered method

different to generating entire events: only a map of the phase-space is learned to distribute random points more efficiently ⇝ can guarantee that physics is left untouched



the quest

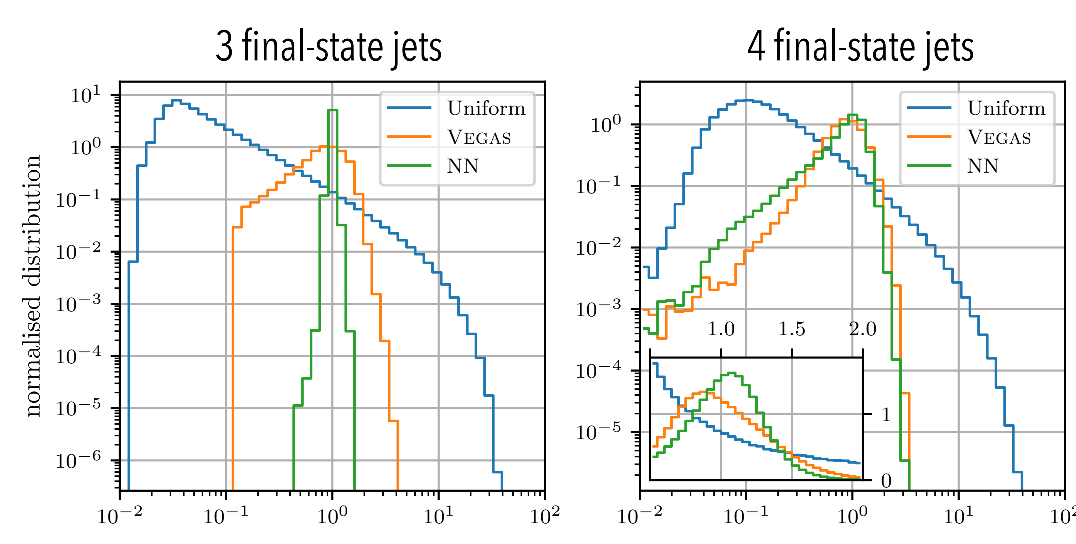
toy examples based on GAN/DNN promising
[Bendavid 1707.00028, Klimek Perelstein 1810.11509]

realistic implementation for physical processes with cuts etc. using …
Normalising Flows [EB et al. 2001.05478, Gao et al. 2001.10028
DNN [Chen Klimek Perelstein 2009.07819]

☺️ promising improvements for simpler cases

☹️ so far not better than VEGAS for more complex cases

⇝ certainly worthwhile to keep trying



[EB et al. 2001.05478]

IMPROVE ME THROUGHPUT

Parallel computing power of CPU vector engines and accelerators under-explored, although MC amenable to parallelisation

previous OpenCL/CUDA attempts (neither reached production quality)

1. MG5_aMC-based: speed-ups ~O(20)-O(200) for various SM procs
   [Hagiwara Kanzaki et al. EPJC 66 477-492] ⤳ picked up again as mentioned by Josh

2. Berends-Giele recursion relations (BG), gg → ng, leading N_C,
   speed ups ~O(150)–O(300) [Giele Stavenga Winter 1002.3446]

time's ripe for a renewed effort at ME@GPU!

10 years ago, trivial parallelisation over CPU cores/clusters was good enough, but now CPU processing capacity has "stalled" and HL-LHC comes closer

(GP)GPU improved: O(10)x memory bandwidth, even more for DP FLOPS

new abstractions address wide scope in parallel computing power and are less vendor-specific: KOKKOS, SYCL, recent OpenMP, OpenACC, HPX etc.

reduction in energy consumption can be substantial
[Tian Benkrid 10.1145/1862648.1862656]

the plan

pick up BG, which has best scaling for high multis

determine speed-ups with Tensorflow, CUDA, SYCL, … and recent GPU
⤳ find best approach in terms of gain and practicality

embark on full SM generalisation and automation, interface with toolchains to make it usable (as SHERPA plug-in or just via LH/HDF5 event output)

EU Strategy for Particle Physics Update 2020:

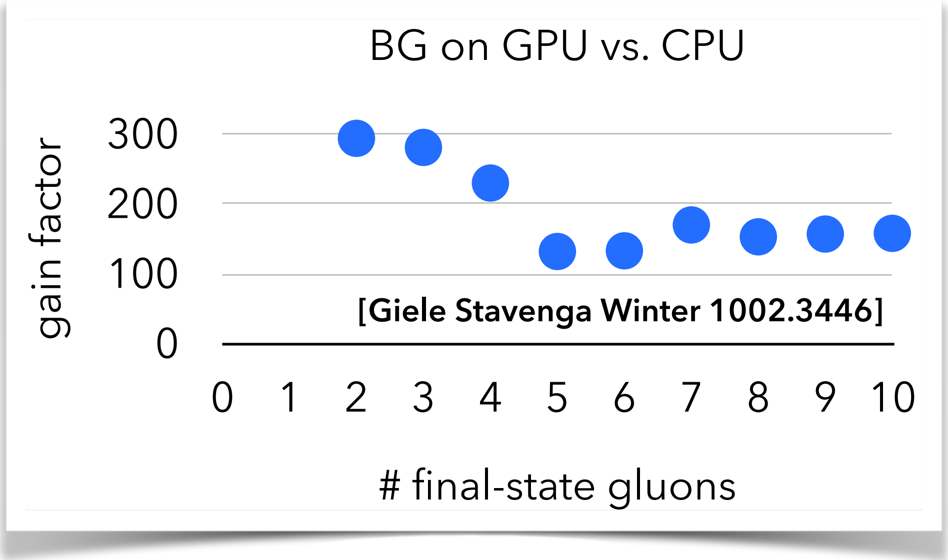
"[d]eveloping accelerator- friendly versions of generators and integrators is an important step in reducing the computing demands for LHC experiments and will be **critical to the success of the [HL-LHC].**"

[1910.11775]

HSF Physics Event Generator WG:

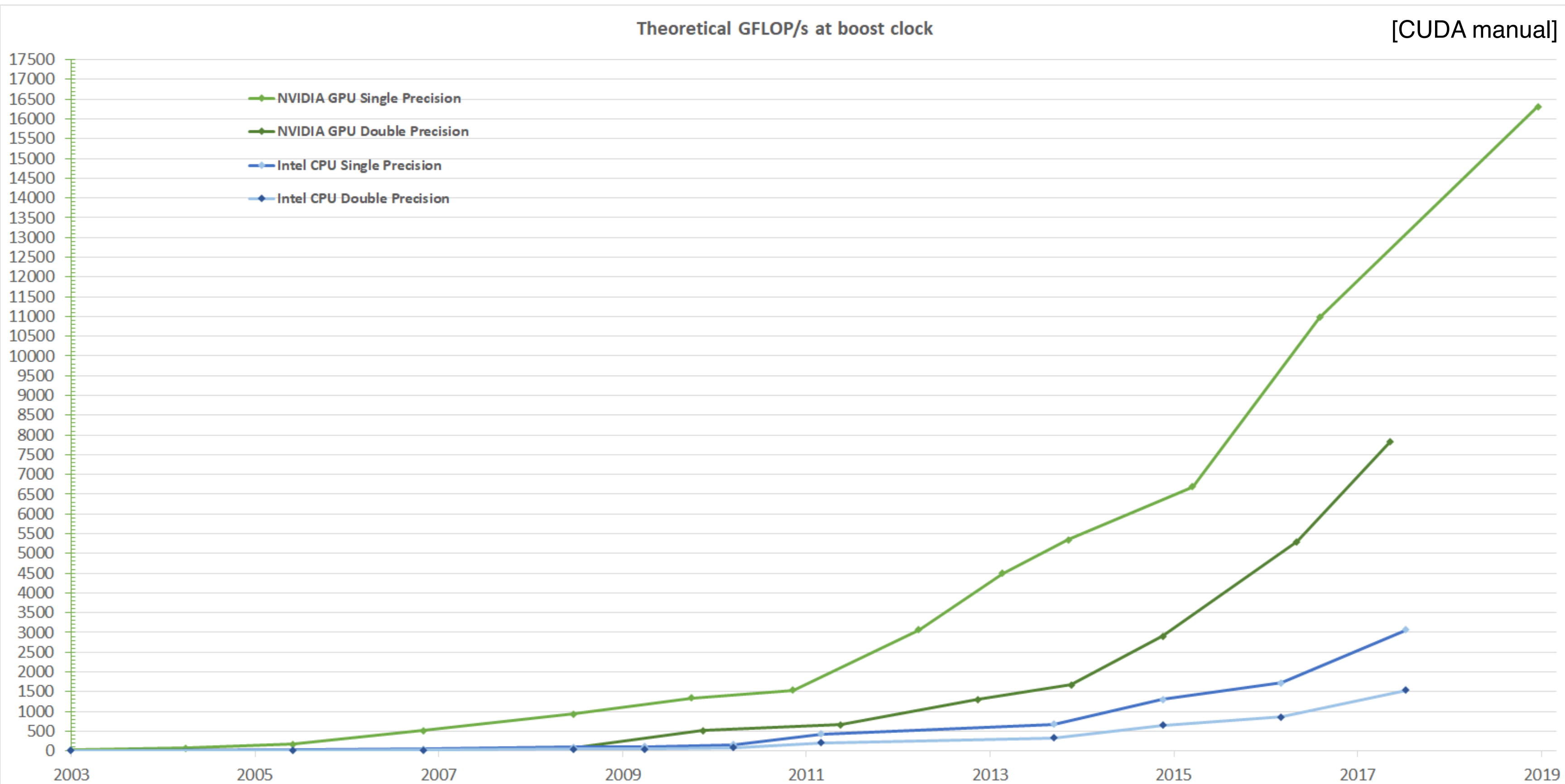
"Porting and optimizing generators on GPUs is especially important to be able to **exploit modern GPU-based HPCs** (such as SUMMIT [145], where 95% of the compute capacity comes from GPUs [146])."

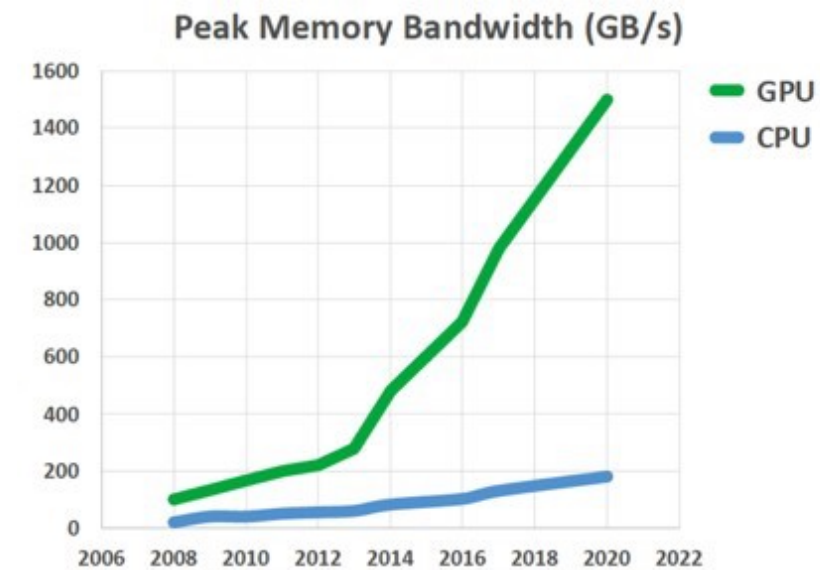
[2004.13687]

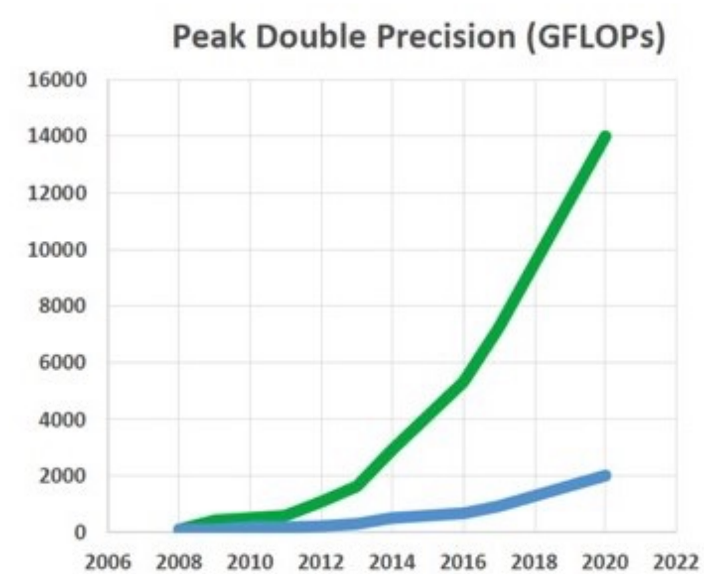


BACKUP

GPU & CPU TREND: DP VS. SP



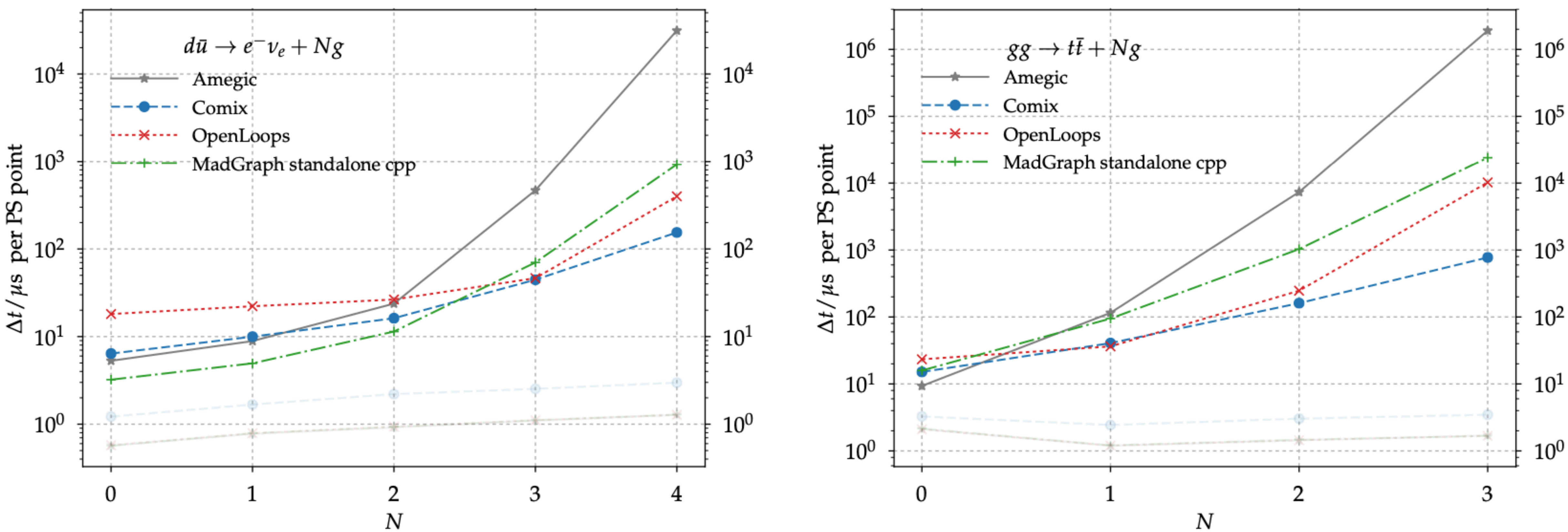




For 2019, the comparison is made between the Nvidia Volta V100 GPU and the Intel Cascade Lake Xeon SP and the trend is projected into 2020.

[https://www.nextplatform.com/2019/07/10/a-decade-of-accelerated-computing-augurs-well-for-gpus/]

TIMING OF INTEGRAND



- Amegic hep-ph/0109036 $\rightarrow$ Feynman diagrams
  Worst case scaling factorial with particle multiplicity
- Comix arXiv:0808.3674 $\rightarrow$ Color-dressed recursion
  Worst case scaling exponential with particle multiplicity
- MadGraph arXiv:1405.0301 $\rightarrow$ Feynman diagrams
  Worst case scaling factorial with particle multiplicity
- OpenLoops arXiv:1907.13071 $\rightarrow$ Color-ordered recursion
  Worst case scaling $\sim$ factorial with particle multiplicity